Draft Notes: OU-7 Status Telephone Conference Call
November 29, 2017

Attendees: Dante Rodriguez (DR)-EPA; Jeryl Gardner(JG)-NDEP; Don Bransford (DB), AECOM; John Roseberry (JR), Brian McCulley (BMc)-Coppers; Ginny Hatch (GH), YPT; Dietrick McGinnis (DM), McGinnis and Associates; Randy Miller (RM), Broadbent; Jamie Tull (JT), ARCADIS; Alma Feldpach (AF), Ramboll; John Batchelder (JB)-Envirosolve; Tashina Jim (TJ), WRPT; Jonathan Shireman (JS), John McMillan (JM), JC Isham (JCI)-APTIM

NOTE: Bryan McCulley is new to Coppers and to the Anaconda project. He comes with 40 years' experience, is a hydrogeologist with extensive experience with mine sites and CERCLA, and is now a Deputy Program Manager for Coppers. One of the strengths he will bring is to develop the statistical method for identifying and discriminating between agricultural and mining impacts signatures in the sample data sets.

The conference calls touched on all the agenda items in the evite, though not in the order of the evite bullet list. Key points by evite bullet item follow:

EPA to send summary of analyses not yet presented.
JB said they do not as yet have a plan for how to address this. DR suggested that DB and Jack Oman talk directly with him to make a plan to address these items.

USGS background info to be compared to BCLs.
DB said that a preliminary review had been done, and a more detailed review was underway. Thanked JS for forwarding the needed reference.

Geochem association of metal; APTIM will confirm if dealt with in statistics call (e.g., agreement on box and whiskers plot).
JS confirmed that they had agreed on a combination of box-and-whisker plots, and major element versus contaminant of concern scatter plots. DB concurred.

Table 4-2, DB's team will do requested highlighting.
DB noted that the table with exceedances highlighted had been produced, and would be forwarded shortly.

Regarding molybdenum and selenium, DB team to review issue (likely next week) and will discuss with JS.
DB said the issue had been resolved with JS, and that there was sufficient data to calculate BCLs specific to OU7 data. JS concurred as long as the OU7 RI does not reference the Area A-1 BCLs.

DM will confer with client and report back on schedule for submitting YPT data report.
DM said they have made some progress. Will be ready to set some deadlines on next week's monthly call. He said the discussion of this call will help him with ideas on how to handle the YPT data.

CSM update document from AECOM team anticipated in early December.

BMc noted that there will be new material in forthcoming CSM document on metal signatures (including Cu and Zn and related constituents). Organic carbon is also being looked at to help distinguish mining versus agricultural influences. Document will include an update of the hydraulic model of the drain, and an update of the CSM figure. He said report was undergoing internal review with issuance to EPA anticipated in December. RM/DB concurred with this synopsis. DB noted that there is no update to the background concentration issues in the forthcoming CSM update as it is still being worked (see below).

APTIM to send agricultural information resources/contacts emails to DM.

Regarding the gathering of agricultural information, BMc confirmed we will be seeing an agricultural timeline in Appendix 1 of the CSM update. It includes information on acres under cultivation and grazing, under fertilizer application, and under irrigation from the late 1800's to present. Also historical information on land use and land cover in Lyon County over time is included. DM asked if agricultural contacts had been sent to him. JM to check.

On statistical topics: 2 sample tests will be done for north/south data, shallow versus deep to be considered by ARC team. There was agreement that total depth BCLs good for comparison to soil/sediment in drain bottom, but prefer to see BCLs by surface and subsurface (and possible north versus south) for deeper and non-drain samples.

This item prompted an extended discussion regarding the calculation of BCLs and handling of the background data. DB opened by saying that before they went further on the depth specific analyses, they wanted agreement on how the results would be used.

He pointed to the issue that splitting surface/subsurface background samples reduces the sample size in each data set, and for a need for EPA to find/acknowledge that the resulting sample sizes are acceptable. Net result would be a sample size is 16 to 22 samples. EPA needs to agree that smaller sample size sufficient. Also, they intend to split the samples for depth into a "0 to 2 ft" data set and a ">2 to 6 ft" data set. Shireman agreed that we do need to look at sample size. If tests for north versus south show that they should be one set, then they can be combined, otherwise analyze separately. Then you can do depth-specific "T tests" for each for set. *(Editorial note; in our conversations use of the term "t-test" is too specific. While this is one type of statistical discriminate test, the general term is "two-sample hypothesis tests" or more succinctly "2-sample tests". This goes back to the original statistics conference call where we discussed first determining the probability distribution and selecting the appropriate tests and parameters to use based on that. In the future, the group should be careful to use general terms where appropriate and specific terms when the discussion needs to be specific).*

DB said they expect the "T tests" to show north and south data sets to be statistically different, then they will need to look at shallow versus deep. DB also noted that there is concern that the smaller sample sizes would have bearing on the outlier assessment. There was a general discussion revolving around the appropriate acceptance criteria for the sample size. JS recommended a 95% confidence in the BCL value, but also suggested using the functionality of ProUCL to determine if the sample size is adequate.

DR asked if they would first look at full depth BCLs, and then look at depth-specific BCLs.

DB said they propose the general breakdown of the depth intervals to be 0 to 2 feet for surface soil and >2 feet to 6 feet for subsurface soils. These soil divisions would serve as the basis for the depth-wise 2-sample hypothesis tests *(often referred to as T-tests in the discussion)*. Similarly, there would be comparisons between the same depth intervals in the North and South background areas. EPA agreed to this in general principal.

Depending on the results of the "T-tests" the data for a specific COI in samples for the depth intervals may be pooled, or may be pooled for a depth interval between the North and South background areas where the tests show there is no statistically significant difference in the sample populations for a particular analyte. If the "T-test" indicated the depth intervals were different (i.e. represented different populations), then depth-specific BCLs would be calculated. A similar approach would be used to establish whether to pool or not pool the data for a specific COI between the North and South areas.

They would use the shallow BCL for shallow samples away from the drain (the deep BCL for the deeper samples). The deep BCL would be used for the ditch bottom samples.

DB said they wanted EPA to agree on the approach ahead of time. Both DM and JS voiced the sentiment that rather than requiring concurrence to a specific approach, ARC should do some preliminary analyses and then bring the results to the group for discussion. DM suggested that the level of effort to establish if the method was viable did not represent a significant level of effort, and that the results should speak for themselves. *(Note: APTIM would suggest that while the group fleshed out the approach and reached some general agreements on the call, the final approach should be communicated in written form by ARC so that the statistical details of the method are clear.*)

DR asked what would be done if we find shallow or deep data sets do not have enough samples. DB confirmed that ARC has no intention of collecting any more background data, and therefore if the shallow or deep dataset is insufficient to calculate a BCL, they will go with a 0 to 6 ft BCL.

BMc suggested he was not sure the shallow versus deep background issue would really make a difference to the risk considerations. AF said if we lack a depth-specific BCL for one or more constituent, they would use a 0 to 6 ft BCL. She was not sure the depth-specific BCLs would be that different, and did not see value in continued delays over the issue.

DB wanted to establish if a path forward had been agreed upon; DR stated his understanding of ARCs proposed approach as follows:

Do "T-tests" on shallow vs deep and north vs south. (*EPA understands that, based on previous agreements, the actual statistical test used will depend on the best-fit probability distribution as indicated by goodness of fit tests on the data using ProUCL or other acceptable statistical analytical methods.*) If for a COI, the datasets are different to 95% confidence level, including depth-specific, or Area-specific, or both comparisons, backgrounds will be calculated for each distinct dataset. If not different, then the full pooled data set will be used to calculate the BCL for that COI. *(A basic flow diagram reflecting this discussion is attached, though it is not a detailed decision tree. APTIM would suggest that ARC provide a brief narrative and decision tree*

diagram on their understanding of how the statistical evaluations are to proceed and the results used.)

Uncertainty relative to background can be discussed in the uncertainties section of the risk assessment, and handled in the risk management decisions. DR suggested that uncertainties be handled with conservativeness. Where a full depth background is used, identify a range in background for that COI. *(Note: the group did not define how a range would be established [e.g., percentiles, upper 95% UTLs, 95% confidence of the mean].)*. Drain bottom samples will be compared to the deep background where one is calculated.

DB said they would send a technical memo with the results

JR to continue working on outlier tests, seems there is agreement.

Per DB, they are on hold on this issue until we have the path forward on background calculations established (see discussion above).

Next monthly check in: Tuesday, January 16, 2018, 10-12 Pacific.